Pointing and Describing in Referential Communication: When Are Pointing Gestures Used to Communicate?[1]

Adrian Bangerter and Eric Chevalley
Institut de Psychologie du Travail et des Organisations
University of Neuchâtel
Rue de la Maladière 23
CH-2000 Neuchâtel, Switzerland
adrian.bangerter@unine.ch

Abstract

What functions, if any, do gestures serve in conversation? Recently, it has been suggested that the function of gesture may vary within a particular type, as a function of the immediate communicative context. Thus, it may not be a question of *whether* a type of gesture is used to communicate, but *when*. We investigated this possibility for the case of pointing gestures in a referential communication task. Pairs worked together to identify targets (photos of people) among pictures in an array that they both could see. They were free to talk and gesture. We manipulated two factors: ambiguity (the number of pictures in an array) and partner visibility. Visible pairs could see each other, hidden pairs could not. We distinguished between full and partial pointing gestures. Full points involve raising the arm whereas partial points do not. Full points were used more often by visible than hidden pairs, thereby suggesting that they are intended to communicate. Their use reduced verbal effort. Their frequency decreased with increasing ambiguity. The use of partial points did not vary with visibility, suggesting that they are automatic in production.

Keywords: Pointing gestures, conversation, referential communication.

1 INTRODUCTION

1.1 THE FUNCTION(S) OF GESTURES IN CONVERSATION

What functions do gestures serve in conversation? Given their ubiquity, one seemingly obvious function is communicating. By this view, gestures are used by the speaker to add information to speech, e.g., by expressing semantic content complementary to verbal utterances. However, a long and diverse strand of research has questioned this possibility. In an influential article, Rimé and Schiaratura (1991) proposed that co-speech gestures are primarily related to speech production, and that their communicative function is peripheral. Using a wide range of methods and research strategies, experimental studies have indeed shown that gestures are related to speech production (Krauss, 1998). One study by Chawla and Krauss (1994) found that gestures are more frequent in spontaneous speech, where production is more demanding, than in rehearsed speech. Another (Morrel-Samuels & Krauss, 1992) demonstrated that gestures concurrent with unfamiliar words are longer in duration than those concurrent with familiar words. Gestures also more generally aid in organizing cognition, as when they facilitate counting activities

[1] This study was supported by a grant (101411-103472) from the Swiss National Science Foundation to the first author.

in preschoolers (Alibali & DiRusso, 1999) or maintain information in spatial working memory (Morsella & Krauss, 2004).

According to these findings, gestures facilitate lexical retrieval, message conceptualization (Kita, 2000), or thinking, in other words they are functional for speakers. To demonstrate a communicative function of gestures in conversation, it is necessary to show (1) that they are designed by speakers for addressees, and (2) that they actually have an impact on comprehension (Krauss, Morrel-Samuels, and Colasante, 1991), in other words the demonstration must focus on activities of both speakers and addressees. Moreover, gestures should be elicited in communicative rather than non-social situations (Alibali, Heath, and Myers, 2001). Indeed, researchers arguing for a communicative function of gesture (Bavelas, 1994; Kendon, 1994) often point out the importance of studying spontaneous gestures produced in natural conversation and their relation to the immediate communicative context.

The possibility that gestures are designed by speakers for addressees constitutes a case of audience design (Clark & Murphy, 1982). Audience design can be shown by comparing the frequency of gestures produced in situations where communicators are visible to each other with the frequency of gestures produced in situations where they are not. If gestures are designed by speakers for addressees, then a lack of visibility should suppress gesture production. Note that, even in this case, the question of the precise level of intentionality of the gesture remains difficult to answer (see however Melinger & Levelt, 2004). If, however, gesture production is insensitive to this change in context, then that is a strong indicator that it is automatic. Gesture frequency does indeed differ according to visibility (Alibali et al., 2001), but not systematically (Rimé, 1982). Sometimes, gestures are produced more frequently in visible conditions but are not entirely suppressed when communicators are not visible to each other. A particularly subtle variation of this technique was used by Özyürek (2002). She found that speakers changed the orientation of their gestures to accommodate to addressees seated at different angles.

Showing that gestures facilitate comprehension is more difficult. Sometimes, gestures are not attended to by addressees, are not well-remembered, or are not easily interpreted independently of co-occurring speech (e.g., Krauss, Morrel-Samuels, & Colasante, 1991; Krauss, Dushay, Chen, & Rauscher, 1995). Early studies (e.g., Graham & Argyle, 1975) have, however, found that gestures can facilitate comprehension in communicative situations. Gestures can even enhance the impact of television advertisements (Beattie & Shovelton, 2005). Other work (Bangerter, 2004; Kelly, Barr, Church, & Lynch, 1999; Louwerse & Bangerter, 2005; Thompson & Massaro, 1986) has shown that pointing gestures affect comprehension.

Of course, the possibility that gestures are related to speech production is not incompatible with the possibility that they are communicative (Alibali et al., 2001; Bavelas, 1994; Özyürek, 2002). These functions are interrelated in other aspects of language use. For example, fillers (e.g., *uh* or *um*) reflect both speech production and interactional processes (Clark & Fox Tree, 2002). Thus, the finding that gestures facilitate speaking does not constitute evidence against the claim that they are communicative, especially when they have been elicited in non-communicative situations (Melinger & Levelt, 2004). Gestures may have both speaker-related (i.e., speech production) and addressee-related (i.e., communicative) functions. In the wake of this conclusion, an emerging trend in the field is the detailed study of different types of gestures and their relation to speech and to the immediate conversational context. Gestures are increasingly viewed as multifunctional and integrated with speech (e.g., Bavelas & Chovil, 2000). Alibali et al. (2001) found that the production of representational gestures (a category subsuming iconic, deictic and metaphoric gestures) varied according to visibility but that the production of beat gestures did not. Nonetheless, representational gestures were also produced in non-visible conditions, indicating that they may serve both cognitive and communicative functions. The authors suggested that research on gesture should "examine how different speakers use gestures in different types of contexts for both speaker-internal and communicative purposes" (p. 186). Bavelas (1994) even goes beyond this in arguing against a purely taxonomic approach to gesture: "In short, the goal of analysis should not be to decide in which category we should put a gesture (or all gestures) but rather to discover at least some of the things a gesture is doing at its particular moment in the conversation." (p. 204).

In sum, it has been proposed that gestures be considered as multifunctional, in other words that different types of gestures be compared, that functional differences may hold even within gesture types, and therefore that they be studied in detail and in relation to their conversational context. Taking into account the immediate context of gesture production may reveal *when* they are communicative and when they are not, rather than *whether* gesture in general is communicative or not or whether a particular type of gesture is communicative or not. The present study takes up these calls. We investigated pointing gestures in a referential communication task. Our goal was to investigate the precise functions of pointing, be they communicative or not. We conducted a controlled experimental study that nonetheless allowed for conversational interaction. Such a setting allows manipulation of key variables (e.g., mutual visibility of interaction partners) while retaining features of naturalistic situations to analyze in detail the relation between gesture and speech. The case of pointing is theoretically interesting because it has received comparatively less attention than other types (for example iconic gestures), and also because pointing seems to convey different information or have different functions than, say, iconic gestures. In the next section we briefly review research on pointing gestures, focusing on the relation between pointing and speech in conversational situations between adults.

1.2 POINTING GESTURES

Pointing gestures are a particular form of deixis. Classical theorists have described their relationship to other types of gestures and to language. Peirce (see Buchler, 1940) classified signs into three categories: icons, indices and symbols. Icons bear a perceptual resemblance to their object of reference (e.g., iconic gestures that mimic a particular manner of movement). Symbols (such as words) are arbitrarily related to their referents. Indices (e.g., pointing gestures) are related to their referents by a physical connection. They are ways of focusing attention. More recently, Clark (2003) proposed that there are two basic kinds of indices, pointing and placing. With pointing, people move the gaze of their addressee towards the referent. With placing, people move the referent into the field of view of the addressee. Pointing and placing thus constitute two ways to focus attention. Another early theorist, Bühler (1965), argued that deictic words primarily serve to direct the attention of the addressee. According to the functional view of gesture advocated by Bavelas (1994), different kinds of gestures (e.g., manual pointing or gaze) can be considered deictic insofar as they serve to focus visual attention. It is even possible to conceive of "chains" of indices. In an intriguing example reported by Marslen-Wilson, Levy, and Tyler (1982), narrators telling a story while pointing to comic-strip characters first looked away from their addressee, down towards their hands. Only then did they point. In other words, they were using their gaze to guide their addressees' gaze to their hands, and then using their hands to point to the referent of their speech.

Several studies have shown that knowing the addressee's current focus of attention is an important resource in conversation. In such situations, speaker and addressee share a joint focus of attention. They are both looking at the same region of visual space, and are aware that the other is also doing so. The number of potential referents is reduced to a subset of all possible referents (Beun & Cremers, 1998; Schmauks, 1991), which allows participants to use reduced verbal descriptions to identify an object. For example, in an experimental task where a director instructs a matcher as to how to build a model from blocks of several colors, an utterance such as "take a red block" will be ambiguous if there are several red blocks among the pieces to be used. But if the subset of blocks within the current joint focus of attention only includes one red block, then this reduced utterance is sufficient. In collaborative physical tasks in mediated settings, having access to information about a partner's gaze (for example by seeing a partner's eye movements superimposed on a shared document) results in more efficient communication (Kraut, Fussell, & Siegel, 2003; Velichkovsky, 1995). Gaze may also facilitate communication by allowing speakers and addressees to monitor each other and repair utterances as they are produced (Clark & Krych, 2004). Of course, analogous processes operate within language, as when focus spaces constrain pronominal reference (Brennan, 1995; Grosz & Sidner, 1986).

These accounts of the function of pointing gestures contrast with a "standard" view (Lyons, 1982) of deixis. According to this view, a pointing gesture serves to identify the referent of a deictic expression such as *that's my car*. Focusing attention does not play any significant role in this account. This assumption is embodied some experimental studies. For example, in a situation where a child has to pick out one present among four possible toys, a pointing gesture is insufficient to discriminate among them given that the toys are all placed too close to each other. In such situations, pointing has been described as less versatile, flexible and effective than language: "Linguistic devices, being the most versatile ones, seem to make other forms superfluous" (Pechmann & Deutsch, 1982, p. 331). Although this may be true for experimental settings where the task is to identify a target within a well-specified referential domain, a more typical function of pointing in naturalistic conversation (and one where it may be more effective) may be precisely to *construct* such a situation by focusing attention on a reduced domain of referents. This line of reasoning is of course derived from the joint-focus-of-attention view explained above.

How can these two accounts be reconciled? When is pointing used to focus attention, thereby indirectly facilitating reference, and when is it used to directly identify or locate a referent? To answer this question, we draw on recent proposals that view referring as a composite signal (Clark, 1996; Clark & Bangerter, 2004; Engle, 1998; see also Bavelas & Chovil, 2000), i.e., as typically consisting of both descriptive and indexical components. This proposal is consistent with observations from field research (Goodwin, 2003). Consider the following example (Schegloff, 1984, p. 280):

(1) Frank: why:nchu put that t the end uh the ta:ble there [pointing]

Here, Frank uses a combination of describing (*t the end uh the ta:ble*) and indicating (*there* accompanied by a pointing gesture) to get his addressee to recognize where to put a dish. What determines the combinations of describing and indicating that are used in referring? An important aspect of the composite signal view is that referring is flexible in both production and comprehension. In other words, speakers are capable of adapting their messages mid-utterance to fit addressees' evolving signals of comprehension (Clark & Krych, 2004). Likewise, addressees are capable of rapidly integrating speaker utterances with other sources of information (Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002). The relative importance of describing and indicating in a referring act can be opportunistically adapted to the situation. In short, composite signals exhibit a tradeoff between describing (typically accomplished linguistically, except in the case of iconic gestures) and indicating (typically accomplished gesturally). Something similar holds for comprehension: People rely differently on linguistic and gestural components in comprehension depending on their relative ambiguity (Thompson & Massaro, 1986£). Thus, the informativeness of a pointing gesture may be limited by its ambiguity. There is an upper limit on the accuracy with which people can detect where another person is pointing (Bangerter & Oppenheimer, 2006). If the situation is ambiguous, pointing (indicating) alone will be insufficient and will need to be augmented by describing the target. People might then use a pointing gesture to direct their addressees' gaze to the target region and subsequently describe the object to pick it out among potential confounding referents. This may especially be the case when the referents of pointing gestures are distant from the pointer (Bangerter, 2004, van der Sluis & Krahmer, 2004). Therefore, we propose that the relative reliance on describing and indicating will vary as a function of the referring situation, and especially as a function of the ambiguity of pointing.

1.3 THE PRESENT STUDY

In the work reported here, we manipulated partner visibility and ambiguity to study their effects on the use of describing and indicating in a dyadic referential communication task. Our goal in manipulating visibility was to understand to what degree pointing gestures are used to communicate. The research summarized above indicates that in order to be considered communicative, gesture must be shown to be sensitive to context (e.g., visibility) in production as well as to have an effect on comprehension (Krauss et al., 1991). In other words, gestures should be produced more frequently and should result in more efficient verbal communication when partners are mutually visible than when they are not. Of particular interest are residual gestures produced in the non-visible condition. Many studies have found that gestures are not entirely suppressed when partners are not mutually visible (Alibali et al, 2001). Rather than using these findings to fuel the debate about whether gestures do or do not communicate, or which types of gestures do or do not communicate, it may be more profitable to analyze variations within types, across situations to see *when* gestures do or do not communicate (Bavelas, 1994). To this end, after an initial examination of the videotapes from our experiment, we decided to distinguish between those pointing gestures where the pointer's elbow was raised off the table and fully extended (hereafter: *full* points) and those where only the forearm was extended, the elbow remaining on the table (hereafter: *partial* points). Full points involve coordinated, purposeful movement of the arm from a resting position in the direction of the target. Thus, it seems likely that they are produced with communicative intent. Partial points do not involve as much movement and may be relatively automatic in production.

Our goal in manipulating ambiguity was to determine how it affects the relative use of describing and pointing in referential communication. When gestures are unambiguous, speakers should rely on them to a larger extent relative to describing. In an extreme case, a referent might be identified simply by pointing to it. When gestures are ambiguous, speakers should rely relatively more on describing a referent to identify it.

2 METHOD

2.1 PARTICIPANTS AND PROCEDURE

Twenty-four French-speaking pairs (director and matcher) worked together. They were seated side-by-side facing a large board supporting an array of portrait photos of people. The array was fully visible to both of them. A plexiglass panel was placed approximately 25 cm from the array to keep participants from touching targets. The task required the director to identify four target photos for the matcher. They talked and/or gestured freely to identify each target. Each target had a name that the director read to the matcher. The matcher wrote the name down on an answer sheet. After all four targets had been identified, the experimenter replaced the array with the next one. They identified targets from 12 arrays.

We manipulated the density of pictures in an array within subjects, and thus the ambiguity of gestural information. There were 2 sets of arrays with 8, 9, 11, 14, 20 and 37 pictures respectively. With four targets, these numbers constitute a 6-point linearly decreasing scale of the average probability of chance identification of a target. For example, the average probability of chance identification is 15.9% in the eight-picture array, 13.6% in the nine-picture array, 10.7% in the eleven-picture array, and so on. The number of pictures defines the density of the array. The denser the array, the closer pictures are to each other, and the more ambiguous gestures will be. Density is our operationalization of ambiguity. Pictures were arranged in a cloud-like fashion in the array, so as not to form obvious rows and columns. The order of presentation of arrays was counterbalanced.

We also manipulated visibility of partners to one another between subjects. In a visible condition (n = 12), pairs could see each other and thus use gestures; in a hidden condition (n = 12), they couldn't. The hidden condition was created by inserting a large wooden board between the director and matcher. The board completely hid them from each other, but did not obscure their view of the arrays.

Thus, the study had a 2 (visibility) by 6 (ambiguity) mixed-model design. The setup is shown in Figure 1.

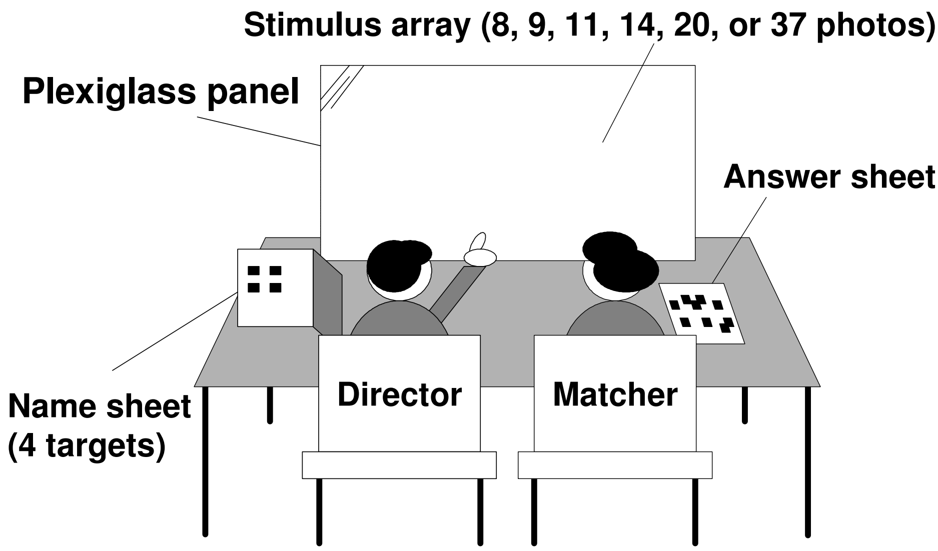

Figure 1: The set-up of the study.

2.2 DATA PREPARATION

Pairs were videotaped with two cameras. Videotapes were mixed onto a split-screen video. Communication was transcribed. Use of gestures and descriptions was coded. Descriptions included absolute spatial descriptions (e.g., *John is on the left*), spatial descriptions relative to another picture (*relative* spatial descriptions, e.g., *Betty is the one below the redhead*), feature descriptions (e.g., *John has glasses*) and deixis (e.g., *right here*, *over there*).

Pointing gestures were defined as a movement of the hand from a resting position (Sacks & Schegloff, 2002) in the direction of the array combined with partial or full extension of one or more fingers. In this respect, they were easy to distinguish from iconic gestures, where the hands were often raised towards the face of the speaker to mimic some feature being described (e.g., downward movement of both hands along the side of the head to mimic long hair). Typical resting positions of pointing gestures were the table or the face (some participants repeatedly touched their faces, especially their mouth or nose). Gesture performance is often decomposed into three phases: preparation, stroke and retraction (Kendon, 1972). The preparation phase involved raising the forearm from the table or removing the hand from the face and moving it in the direction of the array. Depending on the type of gesture, the elbow was sometimes raised off the table. The stroke phase typically involved holding the point for some amount of time (e.g., until an acknowledgment from the matcher) or hand waving, finger extension, flicking or another kind of gesticulation. The retraction phase involved returning the arm to the previous resting position or to another one.

We coded two kinds of pointing gesture, partial and full. Partial pointing gestures were defined as an extension of the arm in which the elbow remained on the table. Full pointing gestures were coded when the elbow was raised off the table. There was a degree of variation in the amount of extension of the upper arm and forearm, typically depending on how much distance participants wanted to cover. Participants sometimes augmented full pointing gestures with fully extended arms by leaning forward in their chairs, especially when pointing at peripheral targets. Examples of partial and full pointing gestures are shown in Figure 2. Both partial and full pointing gestures often involved hand movements at the moment of the stroke.

Inter-rater agreement was assessed by having two independent coders double-code the number of times each type of verbal description and gesture was used per target (irrespective of who used it). Four pairs were double-coded for verbal descriptions and three for gestures. Correlations were computed. They varied between .87 and .93 for verbal descriptions with (all *p*s $< .0001$), indicating excellent agreement. Correlations were .82 for partial points and .92 for full points (both *p*s $< .0001$).

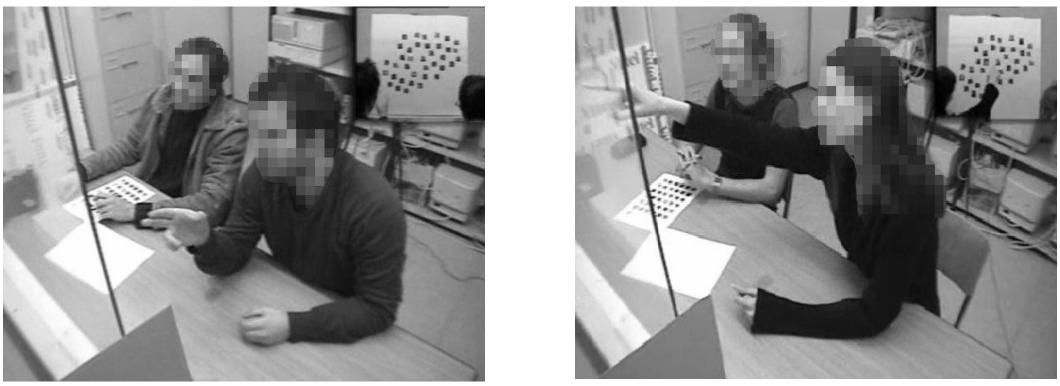

Figure 2: Examples of partial (left side of figure) and full (right side of figure) pointing gestures. The insert in the top right-hand corner of each image depicts an over-the-shoulder view of the task situation.

3 RESULTS

We first present descriptive analyses and examples (3.1). We then analyze the relative use of full and partial points (3.2), verbal effort (3.3) and verbal descriptions (3.4).

3.1 DESCRIPTIVE ANALYSES AND EXAMPLES

Pairs varied in their use of gestures, even within conditions. Some visible pairs used gestures for almost every trial, whereas others did not. Most gestures and descriptions were produced by directors. Identifying targets clearly was easier in the visible condition than in the hidden one. Here is an example of how descriptions were used in the visible condition. Participants are identifying Rachel, on one of the 37-picture arrays (i.e., the highest ambiguity level). Descriptions coded are indicated within parentheses, with the abbreviated type of description in subscript (RLOC = relative location description, ALOC = absolute location description, FEAT = feature description).

1 D Rachel um with (the one next um)$_{RLOC}$
2 M (next to the blonde?)$_{RLOC}$
3 D yeah that's it (with the blue background)$_{RLOC}$

In this example, Rachel is identified with three relative location descriptions and a pointing gesture. The pairs use a presumably more salient adjacent picture, a blonde (2) with a blue background (3), in order to get to her. Below is a pair in the hidden condition trying to identify the same picture, also using a relative location description to begin with.

1 D (she's on your side)$_{ALOC}$ you see (there's a girl with a blue background with her head raised who's laughing the head a bit back ha ha ha)$_{RLOC}$ (1.5) you see?
2 M oh yeah yeah
3 D well (just on the side)$_{RLOC}$ (she's she has her face on the side)$_{FEAT}$ (with a background a bit of red there)$_{FEAT}$
4 M yeah
5 D (and then a smile in the the corner)$_{FEAT}$

Although this pair uses the same strategy, the director and matcher expend more effort in grounding each step of the identification (Clark & Krych, 2004). The director first uses an absolute location description (*she's on your side*) to focus attention on the target region. He then describes the adjacent picture in much more detail (1) before mentioning the target (3, 5). The director also makes sure the matcher is looking at the right picture (*you see*?) before proceeding. In the visible example, there is no feature description, whereas in the hidden example, there are three. In other words, the hidden pair also invests more effort in verifying that the target is the correct one after having tentatively identified it. This is similar to the overspecification observed by van der Sluis and Krahmer (2004) in a production task.

3.2 RATES OF USE OF FULL AND PARTIAL POINTS

We computed the rate of gestures per 100 words (used by both director and matcher) as a dependent measure (Alibali et al., 2001). A higher rate indicates that gestures are used more relatively to words. A lower rate indicates relatively more reliance on words (see Table 1). Visible pairs used full points at a higher rate than hidden pairs. They used less full points per 100 words as ambiguity increased. A 2 (visibility) by 6 (ambiguity) mixed-model ANOVA revealed a main effect of visibility, $F(1,165) = 132.71$, $p < .0001$, a main effect of ambiguity, $F(5,161) = 10.07$, $p < .0001$, and an interaction, $F(5,161) = 10.82$, $p < .0001$. The main effect of ambiguity does not show whether there is an overall increase or decrease, but we tested for a linear trend for visible pairs. It was significantly negative, $F(1,1112) = 15.47$, $p = .0001$. This suggests that full pointing gestures are indeed produced with communicative intent, because visible pairs relied on them more than hidden pairs. It also suggests that pairs relied relatively more on verbal information and less on gestures as ambiguity increased. The rate of partial points per 100 words varied according to ambiguity, but it was not significantly different for visible and hidden pairs: A 2 (visibility) by 6 (ambiguity) mixed-model ANOVA revealed only a main effect of ambiguity, $F(5,165) = 3.28$, $p = .008$. This suggests that partial points are not used to communicate, as their use is not sensitive to visibility.

	Ambiguity (Number of pictures in array)					
	8	9	11	14	20	37
Visible condition						
Full points	7.43 (7.59)	6.34 (6.49)	8.15 (11.9)	5.94 (4.19)	5.32 (3.63)	3.61 (2.41)
Partial points	0.45 (1.80)	1.27 (3.20)	0.55 (2.01)	0.44 (1.96)	0.21 (0.88)	0.40 (1.27)
Absolute location	2.79 (5.17)	5.14 (5.04)	4.47 (5.24)	2.53 (3.16)	2.33 (3.24)	1.73 (2.05)
Relative location	1.51 (4.06)	1.76 (5.05)	1.57 (3.31)	1.73 (4.16)	1.90 (3.77)	2.35 (3.20)
Feature	7.43 (5.29)	5.69 (5.26)	5.47 (4.87)	8.58 (3.73)	7.35 (3.95)	7.16 (4.06)
Deixis	3.88 (5.35)	3.43 (4.59)	3.76 (4.59)	3.87 (3.83)	3.78 (3.83)	2.78 (3.02)
Hidden condition						
Full points	0.04 (0.32)	0.17 (0.79)	0.31 (0.95)	0.00 (0.0)	0.10 (0.53)	0.13 (0.46)
Partial points	0.24 (0.98)	0.54 (1.60)	0.27 (1.02)	0.44 (1.14)	0.26 (0.74)	0.42 (0.95)
Absolute location	3.82 (3.07)	4.24 (3.97)	3.09 (3.27)	2.90 (2.59)	2.51 (2.24)	2.68 (2.10)
Relative location	1.82 (2.82)	2.37 (4.15)	2.10 (2.61)	2.23 (2.59)	1.48 (1.96)	2.43 (2.26)
Feature	7.14 (3.53)	7.28 (4.65)	6.83 (3.56)	7.79 (3.19)	7.48 (2.85)	5.74 (2.52)
Deixis	0.02 (0.17)	0.09 (0.77)	0.00 (0.0)	0.03 (0.23)	0.03 (0.22)	0.03 (0.27)

Table 1: Mean rates per 100 words of full and partial points, absolute and relative location descriptions, feature descriptions and deixis as a function of visibility and ambiguity.

3.3 VERBAL EFFORT

Mutual visibility reduced verbal effort, and ambiguity increased it. A 2 (visibility) by 6 (ambiguity) mixed-model ANOVA with mean number of words per target as dependent variable revealed a main effect of visibility, $F(1,165) = 141.75$, $p < .0001$, a main effect of ambiguity, $F(5,161) = 38.75$, $p < .0001$, and an interaction, $F(5,161) = 6.44$, $p < .0001$. Figure 3 shows the mean levels and error bars, as well as significant differences (identified by repeated-measures contrasts) between consecutive data points of the same line. The fact that visible pairs needed fewer words to complete the task suggests that pointing gestures do indeed reduce verbal effort. However, the reduction in verbal effort could also be due to confounding factors, such as the fact that visible pairs were much more aware of where their partners were looking. Another, more direct, way of testing whether gesture use aids referential communication is by looking at differences between pairs in the same condition. We correlated the total number of gestures of both types with the total number of words for each pair ($n = 12$). Visible pairs that used more full points used less words, $r = -.81$, $p < .001$, but their total number of partial points was unrelated to the total number of words, $r = .33$, *ns*. Thus, the lower verbal effort in the visible condition is at least partly related to gesture use. For hidden pairs ($n = 12$), we found that their total number of full points was not related to their total number of words, $r = -.16$, *ns*. However, their total number of partial points was *positively* related to their total number of words, albeit marginally, $r = .57$, $p = .052$.

3.4 VERBAL DESCRIPTIONS

Because verbal effort varied according to ambiguity and visibility, we computed the mean number of times verbal descriptions (absolute location descriptions, relative location descriptions, feature descriptions) were used per 100 words (irrespective of who used them). This is their rate of use or *relative* use. We also computed the rate of use of deictic expressions (e.g., *here*, *there*). Descriptive data are shown in Table 1. Feature descriptions were most often used, followed by absolute location descriptions and relative location descriptions. Deixis was used regularly only by visible pairs, suggesting that they accompanied pointing gestures. In what follows, we report inferential statistics for each type of verbal description.

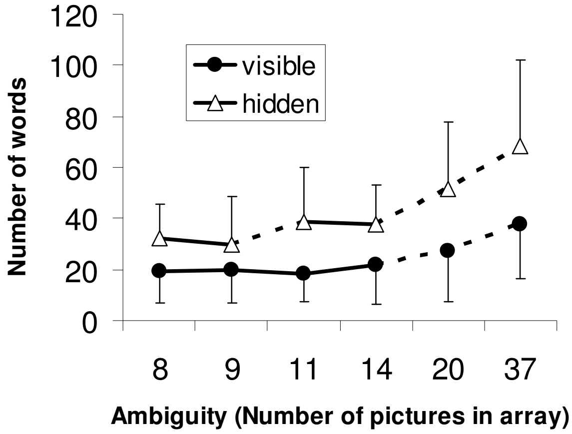

Figure 3: Mean number of words per target as a function of visibility and ambiguity. Error bars indicate one standard deviation. Dotted lines indicate significant differences between consecutive points.

The relative use of absolute location descriptions varied according to ambiguity. A 2 (visibility) by 6 (ambiguity) mixed-model ANOVA revealed a main effect of ambiguity, $F(5,161) = 10.38$, $p < .0001$, and an interaction between ambiguity and visibility, $F(5,161) = 3.11$, $p = .01$. The main effect of visibility did not reach significance, $F(1,165) = .03$, *ns*. We tested for linear and quadratic trends in the relationship between absolute descriptions and ambiguity. There was a significant quadratic trend for visible pairs, $F(2,568) = 14.58$, $p < .0001$, adjusted $R2 = .045$, i.e., absolute location descriptions were used relatively less often at low and high levels of ambiguity than at intermediate levels. For hidden pairs, there was a significantly decreasing linear trend, $F(1,541) = 25.9$, $p < .0001$, adjusted $R2 = .044$.

The relative use of relative location descriptions did not vary significantly according to density or ambiguity, all Fs < 1.6, *ns*.

The relative use of feature descriptions varied according to ambiguity. A 2 (visibility) by 6 (ambiguity) mixed-model ANOVA revealed a main effect of ambiguity, $F(5,161) = 7.50$, $p < .0001$, and an interaction between ambiguity and visibility, $F(5,161) = 4.70$, $p = .0001$. The main effect of visibility did not reach significance, $F(1,165) = 0.12$, *ns*. There was a significant quadratic trend for hidden pairs, $F(2,540) = 4.28$, $p = .014$, i.e., feature descriptions were used relatively less often at low and high levels of ambiguity than at intermediate levels. However, the size of the effect was small, adjusted $R2 = .012$.

The relative use of deictic expressions varied only according to visibility. A 2 (visibility) by 6 (ambiguity) mixed-model ANOVA revealed a main effect of visibility, $F(1,165) = 164.23$, $p < .0001$, with more deictic expressions being used by visible pairs. There was no effect of ambiguity, $F(5,161) = 1.40$, *ns*, nor was there a significant interaction, $F(5,161) = 1.45$, *ns*.

Thus, taken together, the rate of relative use of verbal descriptions only varied substantially for absolute location descriptions at different levels of ambiguity. Both hidden and visible pairs used them less often at high levels of ambiguity. Visible pairs increased their use from lower to intermediate levels of ambiguity.

4 CONCLUSIONS

The most important finding relates to the distinction between two types of pointing gestures, partial and full points. Pairs used more full points when their partners could see them than when they couldn't. This suggests that full points are produced with communicative intent. Pairs that used more full points also used fewer words to complete the experiment, indicating that full points affect comprehension. Finally, the relative reliance on full points decreased with increasing ambiguity. This is consistent with the composite signal view of multimodal communication: people rely differentially on linguistic and gestural components of a multimodal signal in an opportunistic manner.

The use of partial points was not sensitive to visibility, suggesting that their production is not intended to communicate and may be automatic. As was discussed above for other types of gesture, partial points may possibly be functional for the speaker. They may serve as a visual marker. Some directors looked back and forth between their list of targets and the target on the board, possibly to verify whether they had the correct picture before describing it to the matcher. A partial point may have helped them remember where the target was located on the array, especially for dense arrays. Alternatively, partial points may also reflect difficulties in production. The fact that they were marginally positively correlated with the number of words used by hidden pairs is consistent with this interpretation. Hidden speakers that were having difficulty formulating a description may have spontaneously produced a partial point, much in the same way as they spontaneously produced iconic gestures during describing (Morrel-Samuels & Krauss, 1992). Visible speakers in the same situation may have simply extended their arm to transform the partial point into a full point.

The relative use (i.e., per 100 words) of different verbal descriptions did not vary much with ambiguity. This means that pairs did not change their verbal strategies fundamentally as ambiguity increased. Instead, they simply did more of the same thing. The only reliable trend was a curvilinear relationship between ambiguity and use of absolute location descriptions (e.g., descriptions like *in the middle* or *on your side*). At intermediate levels of ambiguity, absolute location descriptions were used more often than at low and at high levels of ambiguity. Although this finding is difficult to interpret, it may be that absolute location descriptions are unnecessary at lower levels of ambiguity because a pointing gesture suffices to focus attention on the approximate region of the target. This is consistent with the finding that only visible pairs exhibited the curvilinear relationship: only visible pairs are in a position to substitute pointing gestures for absolute location descriptions and thus show a lower rate of use at lower levels of ambiguity. At higher levels of ambiguity, absolute location descriptions may be less useful because more pictures may be in the same approximate target region. This limitation on the efficiency of absolute location descriptions at higher ambiguity levels holds for both visible and hidden pairs and is therefore consistent with the observed data.

The data support the idea that gestures within a given type (here, pointing gestures) may have both communicative and speaker-related functions. More generally, the data show that it may be worth exploring how the precise function of a gesture varies according to the communicative context (Alibali et al., 2001, Bavelas, 1994). It is entirely possible that a full point may start out as a partial point. Moreover, we observed (as did van der Sluis & Krahmer, 2004) that many pointing gestures exhibited some kind of lateral hand movement at the stroke. Informal observation suggested that these movements may be similar to iconic gestures in that they may be related to the production of verbal descriptions. A single pointing gesture may therefore have several different functions at different moments in its production. Taxonomic approaches to gesture may indeed be potentially misleading, as suggested by Bavelas (1994). To conclude, we suggest that more research should attend to the possible variations in gesture function within types and in different communicative contexts.

REFERENCES

Alibali, M.W., & Di Russo, A.A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development, 14*, 37-56.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production. Some gestures are meant to be seen. *Journal of Memory and Language, 44*, 169-188.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science, 15*, 415-419.

Bangerter, A. & Oppenheimer, D. M. (2006). Accuracy in detecting referents of pointing gestures unaccompanied by language. *Gesture*, 85-102.

Bavelas, J. B. (1994). Gestures as part of speech: Methodological implications. *Research on Language and Social Interaction, 27*, 201- 221.

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning : An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology, 19*, 163-194.

Beattie, G. W., & Shovelton, H. K. (2005). Why the spontaneous images created by the hands during talk can help make TV advertisements more effective. *British Journal of Psychology, 96*, 21-37.

Beun, R.-J., & Cremers, A. H. M. (1998). Object reference in a shared domain of conversation. *Pragmatics & Cognition, 6*, 121-152.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes, 10*, 137-167.

Buchler, J. (Ed.) (1940). *Philosophical writings of Peirce*. London: Routledge and Kegan Paul.

Bühler, K. (1965). *Sprachtheorie*. [Theory of language.] Stuttgart: G. Fischer.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., and Carlson, G. N. (2002). Circumscribing referential domains in real-time language comprehension. *Journal of Memory and Language, 47*, 30-49.

Chawla, P., & Krauss, R. M. (1994). Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology , 30*, 580-601.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing: where language, culture and cognition meet* (pp. 243-268). Hillsdale, NJ: Lawrence Erlbaum.

Clark, H. H., & Bangerter, A. (2004). Changing ideas about reference. In I. Noveck & D. Sperber (Eds.), *Experimental Pragmatics* (pp.25-49). Basingstoke: Palgrave Macmillan.

Clark, H. H., & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speech. *Cognition* , *84*, 73-111.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language , 50*, 62-81.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J. F. LeNy & W. Kintsch (Eds.), *Language and comprehension* (pp. 287-299). Amsterdam: North-Holland.

Engle, R. A. (1998). Not channels but composite signals: Speech, gesture, diagrams and object demonstrations are integrated in multimodal explanations. *Proceedings of the 20th Annual Conference of the Cognitive Science Society*.

Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: where language, culture and cognition meet* (pp. 217-241). Hillsdale, NJ: Lawrence Erlbaum.

Graham, J. A., & Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. *International Journal of Psychology, 10*, 57-67.

Grosz, B., & Sidner, C. (1986). Attentions, intentions and the structure of discourse. *Computational Linguistics, 12*, 175-204.

Kelly, S. D., Barr, D., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language, 40*, 577-592.

Kendon, A. (1972). Some relationships between body motion and speech. An analysis of an example. In A. Siegman & B. Pope, (Eds.), *Studies in Dyadic Communication* (pp. 177-210). Elmsford, New York: Pergamon Press.

Kendon, A. (1994). Do Gestures Communicate? A Review. *Research on Language and Social Interaction 27*, 175-200.

Kita, S. (Ed.) (2003). *Pointing: where language, culture and cognition meet*. Hillsdale, NJ: Lawrence Erlbaum.

Krauss, R. M. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science, 7*, 54-60.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61*, 743-754.

Krauss, R. M., Dushay, R. A., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology , 31*, 533-552.

Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human-Computer Interaction, 18*, 13-49.

Louwerse, M. M., & Bangerter, A. (2005). Focusing attention with deictic gestures and linguistic expressions. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society.*

Lyons, J. (1981). *Language, meaning and context*. Fontana Paperbacks.

Marslen-Wilson, W., Levy, E., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. J. Jarvella & W. Klein (Eds.), *Speech, place and action. Studies in deixis and related topics*. (pp. 339-378). Chichester: John Wiley.

Melinger, A., & Levelt, W. (2004). Gesture and the communicative intention of the speaker. *Gesture, 4*, 119-141.

Morrel-Samuels, P., & Krauss, R.M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 615-623.

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *American Journal of Psychology, 117*, 411-424.

Özyürek, A. (2002). Do speakers design their co-speech gestures for their addresees? The effects of addressee location on representational gestures, *Journal of Memory and Language, 46*, 688-704.

Pechmann, T., & Deutsch, W. (1982). The development of verbal and nonverbal devices for reference. *Journal of Experimental Child Psychology, 34*, 330-341.

Rimé, B. (1982). The elimination of visible behavior from social interactions: effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology, 12*, 113-129.

Rimé, B., & Schiaratura, L. (1991). Gesture and speech. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of Nonverbal Behavior* (pp. 239-281). Cambridge: Cambridge University Press.

Schegloff, E.A. (1984). On some gestures' relation to talk. In J.M. Atkinson & E. J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 266-296). Cambridge: Cambridge University Press.

Schmauks, D. (1991). *Deixis in der Mensch-Maschine-Interaktion*. [Deixis in human-computer interaction.] Tübingen: Max Niemeyer.

Thompson, L. A., & Massaro, D. W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology, 42*, 144-168.

Van der Sluis, I., & Krahmer, E. (2004). The influence of target size and distance on the production of speech and gesture in multimodal referring expressions. *Proceedings of the ICSLP-2004*.

Velichkovsky, B. M. (1995). Communicating attention. Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition, 3*, 199-224.